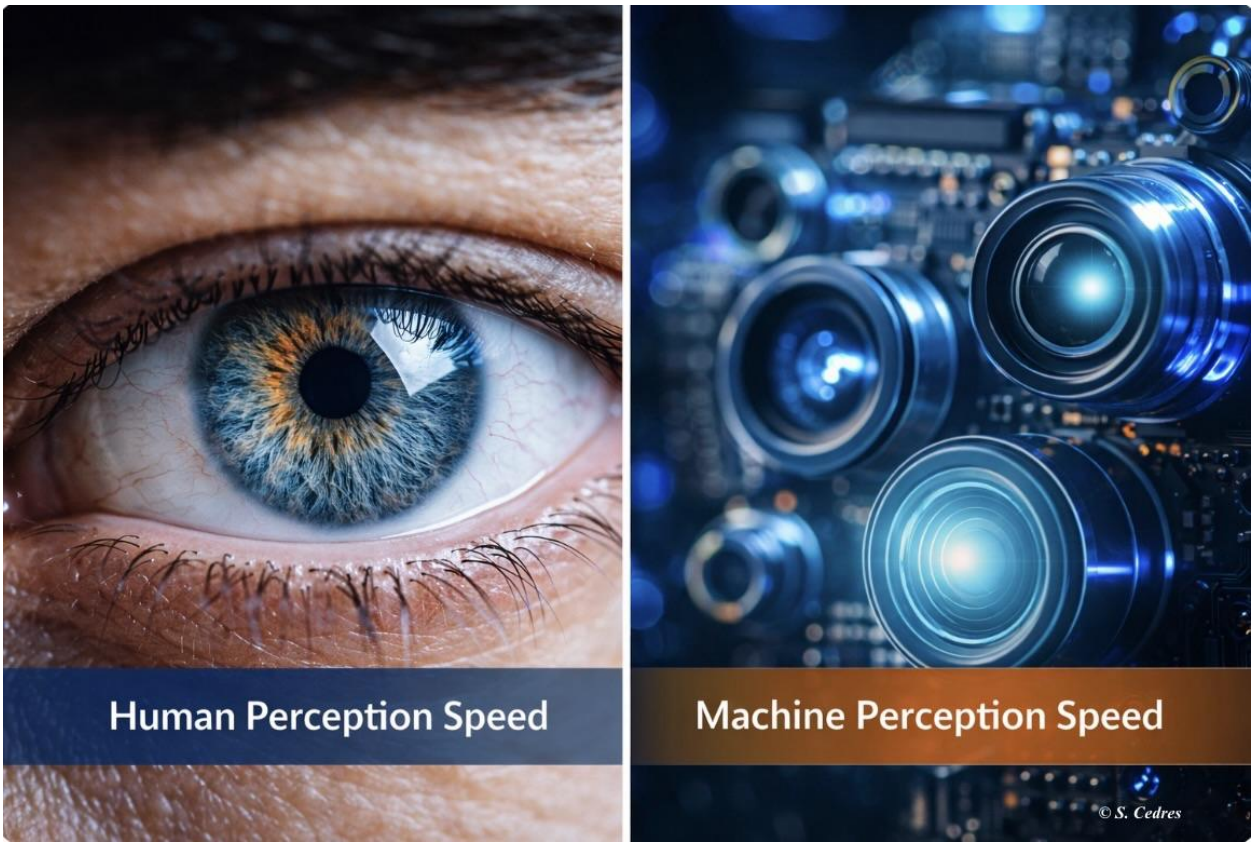

From Human-in-the-Loop to Human-on-the-Loop: Why Latency Is Now the Real Risk

By Stewart Cedres

(March 10, 2026)

Originally published on LinkedIn

For decades, the gold standard for safety-critical systems was the **Human-in-the-Loop (HITL)**. We operated under the assumption that a human should be the final gatekeeper for every significant action.

That assumption made perfect sense when machines were blind, brittle, and context-poor. But today, the environment has shifted, not philosophically, but **physically**.

In the modern grid and industrial landscape, the bottleneck is no longer the machine's intelligence. It is the human's reaction time.

The Physics of the Problem: Latency as Vulnerability

Modern infrastructure operates at timescales that the human nervous system simply cannot match.

Consider the telemetry we now manage:

- **Microseconds:** The speed of electrical faults and circuit-breaker coordination.
- **Milliseconds:** The window for collision avoidance in swarms or detecting thermal runaway in high-density battery arrays.
- **Seconds:** The average human reaction time to a complex, multi-variable alert.

In a high-speed, high-bandwidth environment, **latency is no longer an inconvenience. It is a systemic vulnerability.**

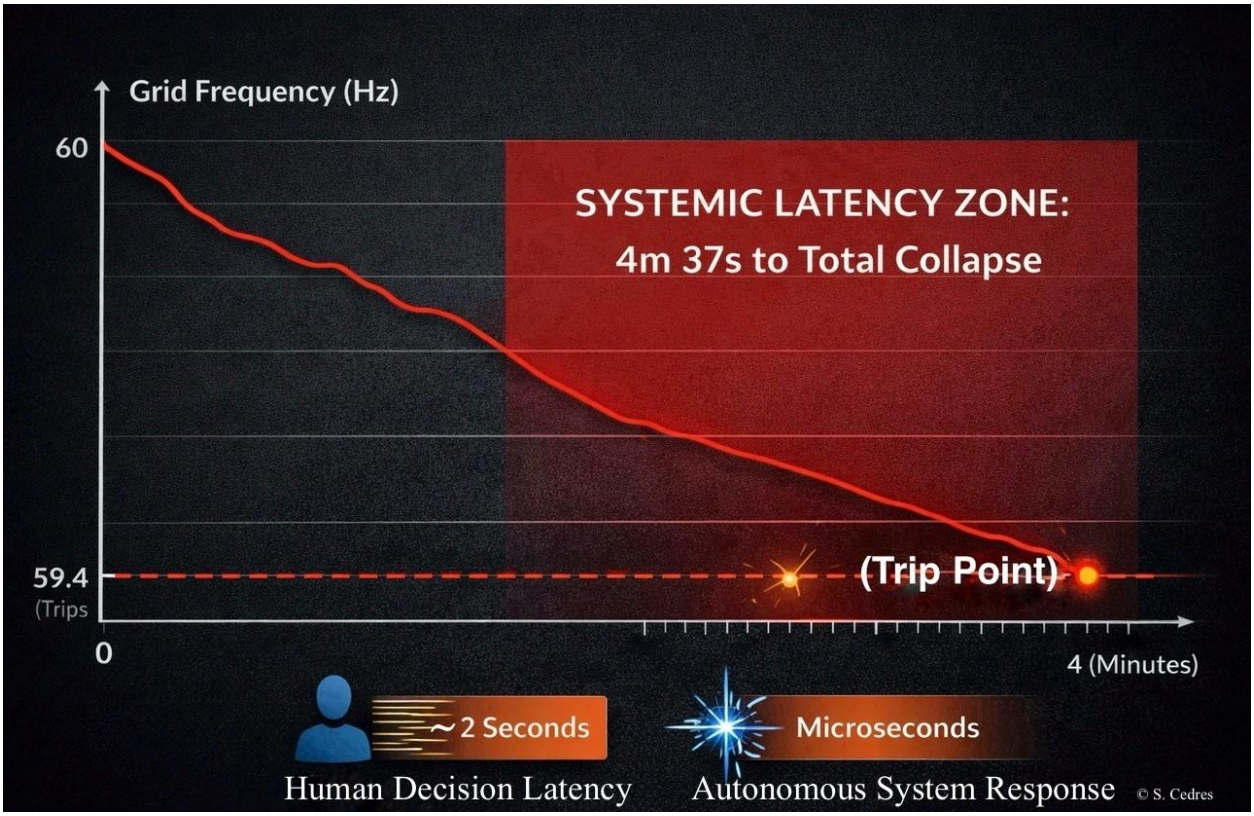

Figure: During the 2021 Texas Freeze, grid frequency approached the 59.4 Hz collapse threshold within a 4-minute 37-second window - far faster than human decision cycles. This is the physics that forces Human-on-the-Loop architectures.

We saw the cost of this vulnerability during the **2021 Texas Freeze**. ERCOT wasn't just fighting a storm. They were fighting a **4-minute and 37-second clock**. As frequency plummeted toward a catastrophic 59.4 Hz, generators began tripping offline automatically

to protect themselves. The human operators were heroes, but they were trapped in an architecture where 'sensing' and 'acting' required manual coordination while the physics of the grid were failing in milliseconds. They weren't fighting against bad intent. They were fighting against **Systemic Latency**.

In the end, the operators saved the day — but the architecture nearly lost the war.

The Scaling Trap

Traditional "Human-in-the-Loop" architectures scale negatively. As we add more sensors, more data, and more complexity, we don't make the human smarter. We make the bottleneck narrower.

When the flow of information exceeds the capacity of the processor—in this case, the human brain—the system doesn't just slow down; it breaks down.

This failure occurs at two critical friction points:

- **Centralized Approval (The Architectural Friction):** This forces a "Global Pause" for a "Local Event." In a networked system, requiring a central human authority to validate a millisecond correction introduces a propagation delay that often exceeds the system's safety margin.

  It treats a high-speed kinetic event like a bureaucratic filing.

- **Cognitive Overload (The Biological Friction):** More data actually slows the decision loop. As the volume of telemetry increases, the operator can no longer tell what matters. The operator isn't just fighting the crisis; they are fighting the sensory saturation caused by the very tools meant to help them.

As a systems engineer, I see this as a classic impedance mismatch. We are trying to control a gigahertz reality with a hertz-speed processor.

The Shift: Human-on-the-Loop (HOTL)

The shift to **Human-on-the-Loop** is not about removing human authority. It is about **elevating** it. It is the transition from micromanagement to **Mission Supervision.**

In a HOTL framework:

- **The Human** sets the rules of engagement, ethical constraints, and safety boundaries.

- **The Machine's** AI-driven nodes handle perception, local decision-making, and immediate corrective action.

The machine operates *within* the human's intent, but at the *speed* of the event. This isn't speculation. It's already the reality in our most high-consequence environments.

While regulatory frameworks are still catching up, the debate is largely settled **by the physics of our most high-consequence environments.** We already trust autonomy to manage the 'un-flyable' surfaces of modern fighter jets and the microsecond load-balancing of high-frequency energy markets. In those sectors, the decision wasn't driven by cost. It was driven by survival — the system demanded a response time that biology simply cannot provide.

Why This Is Inevitable

System speed, environmental volatility, and infrastructure interdependence are all increasing simultaneously. Human-in-the-loop architectures cannot keep up without becoming the trigger for the very failures they were meant to prevent.

This structural correction is the foundation for what I call Autonomous Dynamic Machines (ADM): systems where intelligence and action are coupled tightly enough to eliminate fatal latency.

The Closing Thought

Autonomy is not a moral shortcut or a way to "cut costs." It is a structural necessity for a world moving faster than human reflexes.

The question is no longer *if* humans stay in the loop. The question is **where they belong.** We belong at the level of strategy and intent, leaving the milliseconds to the machines.

The era of the human bottleneck is over. The era of the human supervisor has begun.

Next: *The Bio-Inspired Swarm – Solving for Systemic Friction*

About the Series

This is **Article 1 of 3** in the series: **The Post-Manual Horizon: Engineering the Era of Autonomy.**

- **Article 1: From Human-in-the-Loop to Human-on-the-Loop**
- **Article 2: The Bio-Inspired Swarm: Solving for Systemic Friction**
- **Article 3: Nano-ADM: When Machines Predict and Operate at the Quantum Scale**

About the Author:

Stewart Cedres is a systems engineer with over three decades of experience working at the intersection of autonomous machines, grid resilience, and critical infrastructure. He has proposed the Cedres Integral of Autonomy, a framework describing the structural transition from human-in-the-loop control toward architectures that remove latency as a systemic risk.

Source: Cedres, S. (2026). From Human-in-the-Loop to Human-on-the-Loop. https://www.linkedin.com/pulse/from-human-in-the-loop-human-on-the-loop-why-latency-now-cedres-ren1e